

Frequency Domain capture of light fields using Heterodyning: Analysis of Aliasing

Ashok Veeraraghavan
University of Maryland
College Park, MD 20742
vashok@umiacs.umd.edu

Ramesh Raskar, Amit Agrawal
Mitsubishi Electric Research Labs
Cambridge, MA 02139
{raskar/agrawal}@merl.com

Ankit Mohan, Jack Tumblin
Northwestern University
Evanston, IL 60208
{ankit/jet}@cs.northwestern.edu

Abstract

We analyze the modulation of a light-field via non-refracting attenuators. The design first proposed in [8] exploits frequency-domain modulation to achieve a more efficient encoding. We study the aliasing in the sensed light field and discuss the effects of aliasing on mask based heterodyning light field camera.

1. Introduction

Light fields, characterize the irradiance of each ray in space using a 4 D twin plane parameterization [5][4]. By capturing a light field of the scene, all information content about the scene appearance can be obtained. We present a class of methods using just masks to capture the information content in the light-field. The key idea is to individually attenuate each ray so that appropriate linear combinations measured by the sensor can be used to recover informative parts of the light field.

1.1. Related Work

Light Field Capture: Sensors are limited to be two dimensional surfaces while the light-field that needs to be sensed is 4 dimensional. Therefore, it is necessary to modulate/transform it so that the information in the

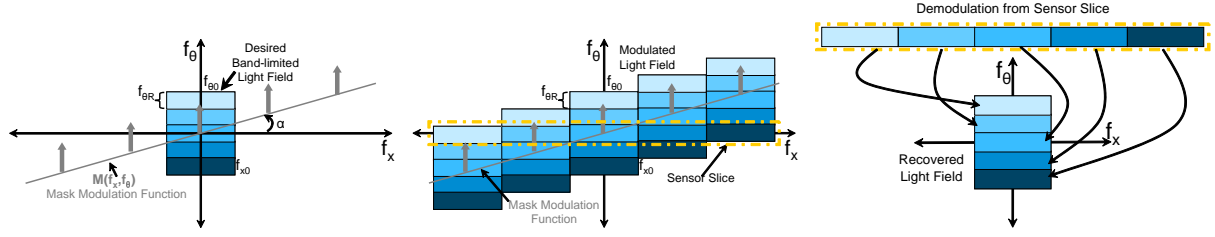


Figure 1. Spectral slicing in heterodyne light field camera. (Left) In Fourier domain, the sensor measures the spectrum only along the horizontal axis ($f_\theta = 0$). Without a mask, sensor can't capture the entire 2D light field spectrum (in blue). Mask spectrum (gray) forms an impulse train tilted by the angle α . (Middle) By the modulation theorem, the sensor light field and mask spectra convolve to form spectral replicas, placing light field spectral slices along sensor's broad $f_\theta = 0$ plane. (Right) To re-assemble the light field spectrum, translate segments of sensor spectra back to their original f_x, f_θ locations.

angular dimensions can be sampled by the sensor. Several optical elements perform this modulation in previously proposed capture devices. A straightforward way to sample angular dimensions is viewpoint sampling. This can be achieved by using a dense array of cameras, one for each viewpoint as in [9]. Such dense camera arrays, however, are impractical for consumer applications because of their sheer bulk. Recently, two handheld light field cameras have been proposed by modifying a traditional camera. The first design uses an array of positive lenses with appropriate prisms in front of a single lens sensor system [3]. The second design (inspired by [2]) uses a microlens array in front of the sensor surface focusing the image of the main lens on the sensor [7]. The array of lenses or the microlens array act as the modulators of the incoming light field in order to enable its capture on a sensor surface. These and other similar devices like beam splitters, mirrors etc. are all refractive in nature, i.e., they bend the incident light wave. But all refractive modulators suffer from inherent limitations such as spherical, chromatic aberrations, coma and misalignment issues.

2. Heterodyne Light Field Camera

To alleviate the light-loss problem of a pinhole array based light-field camera, let us consider a modulator whose frequency response is composed of 5 impulses arranged on a slanted line as shown in Figure 1. This type of modulation was first proposed in [8]. In the primal domain this corresponds to a sum of DC term and two cosines. Therefore, the result of this convolution will be 5 spectral replicas of the light field along the slanted line and therefore the loss of energy is minimal. Moreover, the horizontal slice (red line) of the modulated light field spectrum now captures all the information in the original light field. The RMF for this design is given by

$$R(f_x, f_\theta, :) = \sum_{i=-p}^{i=p} \delta(f_x - if_{x0}, f_\theta - if_{x0} \tan(\alpha), :) \quad (1)$$

Demodulation is done in software by redistributing the 1D signal to the 2D light field space. The process of demodulation consists of rearranging the frequency response of the sensor, $Y(f_s)$, to recover the bandlimited light field $L(f_x, f_\theta)$ as shown in Figure 1. The example was shown for capturing a 2-D light field using a 1D sensor. The very same method can be easily extended to capture a 4D light field using a 2D sensor by using p^2 cosines creating all harmonics of both spatial frequencies f_{x0} and f_{y0} .

Mask based Realization of RMF: As shown above, we need an optical modulator given by $R(f_x, f_\theta, :) = \sum_{i=-p}^{i=p} \delta(f_x - if_{x0}, f_\theta - if_{x0} \tan(\alpha), :)$. This corresponds to an optical modulator, whose energy in the frequency domain is all concentrated on a line in the 2-D space. We already know that when a patterned mask is placed in the path of the incoming light field, in the Fourier domain the effect of the mask is restricted to a line in the 2D space. The required modulation may be achieved by placing a patterned mask whose frequency response is that of a sum of delta functions given by $M(f) = \sum_{k=-p}^{k=p} \delta(f - kf_0)$ and as shown in the Figure 1. This corresponds to a sum of cosines mask with DC and p harmonics of the fundamental frequency f_0 . Moreover if the number of cosines in the mask is given by p , then we also see that the slant angle α should be given by

$$\alpha = \arctan \frac{f_{\theta R}}{2f_{x0}} = \frac{d}{(D+d)} \frac{\pi}{2}, \quad (2)$$

where, d is the distance between sensor and mask while D is the distance between lens and mask.

Solving for the 4D Light Field: To recover the 4D light field from the sensor image, we compute the Fourier transform of the sensor image, *reshape* the 2D Fourier transform into 4D and compute the 4D inverse Fourier transform (Illustration for 2D light field shown in Figure 1(c)). Thus,

$$l(x, \theta) = \text{IFT}(\text{reshape}(\text{FT}(y(s))))), \quad (3)$$

where FT and IFT denote the Fourier and inverse Fourier transforms and $y(s)$ is the observed sensor image.

3. Aliasing

3.1. Data and Aliasing Terms

Let $L(f_x, f_\theta)$ be the incoming light-field band-limited to $(f_{x0}, f_{\theta0})$ respectively. The ray modulation function is given by $R(f_x, f_\theta) = \sum_{i=-p}^{i=p} \delta(f_x - if_{x0}, f_\theta - if_{x0} \tan(\alpha))$ as shown in Figure 1. The Light field after modulation $L_R(f_x, f_\theta)$ is given by,

$$L_R(f_x, f_\theta) = L(f_x, f_\theta) \otimes R(f_x, f_\theta) \quad (4)$$

$$= L(f_x, f_\theta) \otimes \sum_{i=-n}^{i=n} \delta(f_x - if_{x0}, f_\theta - if_{x0} \tan(\alpha)) \quad (5)$$

$$= \sum_{i=-p}^{i=p} L(f_x - if_{x0}, f_\theta - if_{x0} \tan(\alpha).) \quad (6)$$

We can write L_R as, $L_R = L_{Data} + L_{Alias}$ where,

$$L_{Data}(f_x, f_\theta) = L(f_x - kf_{x0}, f_\theta - kf_{x0} \tan(\alpha)) \quad (7)$$

$$L_{Alias}(f_x, f_\theta) = \sum_{i=-p, i \neq k}^{i=p, i \neq k} L(f_x - if_{x0}, f_\theta - if_{x0} \tan(\alpha)), \quad (8)$$

where, k is an integer given by $(2k - 1)f_{x0} \leq f_x \leq (2k + 1)f_{x0}$. When the incoming light field L is bandlimited to $(f_{x0}, f_{\theta0})$, the aliasing terms are all zero, thereby leaving only the data term L_{Data} (refer Figure 1(b)).

3.2. Bandlimit Assumption: Angular frequencies

Let us assume that the band-limit assumption is true in the spatial dimension but is not true for the angular dimension of the light field. This would mean that in Figure 1, the f_x band-limit is true, while the rectangle is actually much taller than is shown in the figure. In this case, the spectral copies created by the $2p + 1$ impulses do not overlap. Therefore, this does not cause mixing between the frequency components and consequently, there is no aliasing. Nevertheless, when the angular band-limiting assumption is not valid, we notice that only the angular frequencies upto the band-limit $f_{\theta0}$ are recovered. This causes some smoothing of the captured light field in the θ direction. Note that there is still no smoothing in the spatial dimension of the captured light field.

3.3. Bandlimit in the Spatial dimension

Traditionally, same-channel masquerading of higher frequencies as lower frequencies due to undersampling, is called aliasing and this usually leads to visually obtrusive artifacts like ghosting. In our camera, when the band-limit assumption is not valid in the spatial dimension, the energy in the higher spatial frequencies of the light field masquerade as energy in the lower angular dimensions. No purely spatial frequency leak to other purely spatial frequency. Thus we dont see familiar jaggies, moire-like low-frequency additions and/or blocky-ness in results. Moreover, we also show using the statistics of natural images that the energy in the aliasing components is very small and therefore the effect of aliasing is imperceptible in real scenes.

Natural Image Statistics: There have been several studies characterizing the power spectrum of natural images that indicate that the energy in the Fourier domain representation of natural images is concentrated in the lower frequencies and this energy decreases atleast at the rate of $1/f$ for higher frequencies [1]. If we use this as a model for natural images, we can show that the energy in the aliasing terms is insignificant. In Equations 78, let us define ϵ_{Alias} as the fraction of each aliasing term to the data term L_{Data} , i.e.,

$$\epsilon_{Alias} = \frac{L(f_x - if_{x0}, f_\theta - if_{x0} \tan(\alpha))}{L(f_x - kf_{x0}, f_\theta - kf_{x0} \tan(\alpha))} \quad (9)$$

Since k is an integer given by $(2k - 1)f_{x0} \leq f_x \leq (2k + 1)f_{x0}$ and $i \neq k$, we know the following about the

arguments in the numerator and the denominator,

$$\begin{aligned} |f_\theta - kf_{x0} \tan(\alpha)| &\leq f_{\theta 0} \leq |f_\theta - if_{x0} \tan(\alpha)| \\ |f_x - kf_{x0}| &\leq f_{x0} \leq |f_x - if_{x0}| \end{aligned} \quad (10)$$

Let us consider $\epsilon_{Alias}(f_x, f_\theta)$ both for values of f_x which are very close to zero and for values of f_x which are close to f_{x0} . For $f_x = f_1 \rightarrow 0$, we see that,

$$\epsilon_{Alias}|_{f_1 \rightarrow 0} = \lim_{f_1 \rightarrow 0} \frac{L(f_1 + (i - k)f_{x0}, f_\theta)}{L(f_1, f_\theta)} \rightarrow 0 \quad (11)$$

When $f_1 \rightarrow 0$, the argument of the numerator is much larger than that of the denominator. Since, the frequency roll-off of natural scenes is at-least $1/f$, ϵ_{Alias} tends to zero. Thus the aliasing term is insignificant. For values of f_x close to f_{x0} , we see that the aliasing term could turn out to be significant but this effect can be suppressed significantly by appropriate pre-filtering or post-filtering.

3.4. Anti-Alias Pre-filtering and Post-filtering

To completely avoid aliasing, we could, in principle, design an optical anti-alias pre-filter that first smoothens the incoming lightfield and ensures that the band-limit assumption is valid. Initially, we considered placing a diffuser near the lens to perform this optical smoothing. Unfortunately, a diffuser placed at the lens will reduce high-frequency angular variations; an unwanted effect. We are most interested in suppressing the high frequencies in the spatial dimension and we were unable to design any optical element that could achieve the desired smoothing effect.

Instead, we designed an appropriate post-filter in order to suppress the effects of aliasing. From Section 3.3, we follow that the effects of aliasing could be significant only near the edges of the band i.e., $f_x \approx f_{x0}$. Therefore, to combat the effects of aliasing near the edges of the band-limit, we filter the recovered light field using a Kaiser-Bessel filter with a filter width of 1.5. Kaiser-Bessel filters have been found to do appropriate post-filtering to suppress the effects of aliasing [6].

4. Discussions and Conclusions

Limitations: Mask based schemes for sensing elements of the light-field are not without their limitations. Any mask based scheme leads to loss of light energy and consequent lower SNR. When the light-field is not band-limited, this might result in aliasing. For scenes with significant energy in the high frequency band above the band-limit assumed, the reconstructions will suffer from aliasing artifacts. We still have not been able to design an optical pre-filter that can perform the required smoothing of the incoming light-field in order to prevent aliasing.

As ray attenuators become finer and finer, they will introduce more and more diffraction artifacts. This might be a limiting factor in terms of the maximum resolution of light fields that may be sensed using ray attenuators.

Future Work: The range of imaging functionalities that may be obtained using a non-refracting attenuator may be significantly increased by considering ray-filters that are controllable both in wavelength and time. By using alternating ray-filters, one for capturing a high resolution image and another for capturing a band-limited light field we may be able to synthesize high resolution light fields. We can also use these high resolution light fields to obtain high resolution texture mapped 3D surface models. Another area of potential future research is novel designs in order to realize highly selective and specific ray-filters using LCD screens and microlens arrays.

References

- [1] R. Balboa and N. Grzywacz. Power spectra and distribution of contrasts of natural images from different habitats. *Vision Research*, pages 2527–2537, 2003. [4](#)
- [2] E.H. Adelson and J.Y.A. Wang. Single lens stereo with plenoptic camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14, 1992. [2](#)
- [3] T. Georgiev, C. Zheng, S. Nayar, B. Curless, D. Salasin, and C. Intwala. Spatio-angular resolution trade-offs in integral photography. *EGSR*, 2006. [2](#)
- [4] S. Gortler, R. Grzeszczuk, R. Szeliski, and M. Cohen. The lumigraph. In *SIGGRAPH*, pages 43–54, 1996. [1](#)
- [5] M. Levoy and P. Hanrahan. Light field rendering. *SIGGRAPH*, pages 31–42, 1996. [1](#)
- [6] R. Ng. Fourier slice photography. *ACM transactions on Graphics*, 2005. [5](#)
- [7] R. Ng, M. Levoy, M. Bredif, G. Duval, M. Harowitz, and P. Hanrahan. Light field photography with a hand-held plenoptic camera. *Stanford University Computer Science Tech Report*, 02, 2005. [2](#)
- [8] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin. Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *Proceedings of ACM SIGGRAPH 2007*. [1](#), [2](#)
- [9] B. Wilburn, N. Joshi, V. Vaish, E. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy. High performance imaging using large camera arrays. *ACM Trans. Graph.*, 24(3):765–776, 2005. [2](#)